Pakpahan E, Hoffmann R, Kröger H.

Retrospective life course data from European countries on how early life experiences determine health in old age and possible mid-life mediators.

***Data in Brief* 2017, 10, 277-282.**

Data in Brief 10 (2017) 277–282

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Retrospective life course data from European countries on how early life experiences determine health in old age and possible mid-life mediators

Eduwin Pakpahan*, Rasmus Hoffmann, Hannes Kröger

Comparative Life Course and Inequality Research Centre, Department of Political and Social Sciences, European University Institute, Via dei Roccettini 9, 50014 San Domenico di Fiesole, Italy

ARTICLE INFO

Article history:
Received 17 October 2016
Received in revised form
21 November 2016
Accepted 30 November 2016
Available online 7 December 2016

Keywords:
Socioeconomic status
Childhood
Old age health
Mediators
Europe

ABSTRACT

The data presented in this article is related to the research paper entitled "The long arm of childhood circumstances on health in old age: Evidence from SHARELIFE" (E. Pakpahan, R. Hoffmann, H. Kröger, 2016) [1]. It presents the distribution of socioeconomic status (SES) and health from childhood until old age in thirteen European countries. In order to capture the characteristics of longitudinal data, which resembles life course data, we divide the data into three schematic periods: childhood (up to 15 years old), adulthood (30 to 60 years old), and old age (61 to 90 years old). This data set contains respondents' life histories, ranging from childhood conditions (such as housing and health) to detailed questions on education, adult SES (working history, income, and wealth) and old age health. The data can be used not only to understand on how early life experiences determine health in old age, but also to recognise the importance of possible mid-life mediators.

DOI of original article: http://dx.doi.org/10.1016/j.alcr.2016.10.003
* Corresponding author.
E-mail address: eduwin.pakpahan@eui.eu (E. Pakpahan).

http://dx.doi.org/10.1016/j.dib.2016.11.094

Specification Table

Subject area	*Ageing, Sociology and Public Health*
More specific subject area	*Life course, longitudinal data, health inequality*
Type of data	*Table and graph*
How data was acquired	*Data was available from SHARE database through registration.*
Data format	*Aggregated, analyzed, filtered*
Experimental factors	*Our sample is based on the SHARELIFE dataset and was extracted using STATA and reorganized using the Stata newspell package* [2].
Experimental features	*For each person, data on health and SES was aggregated for childhood, adulthood, and old age*
Data source location	*Germany*
Data accessibility	*The data was available from the SHARE database through official registration at* http://www.share-project.org/data-access-documentation/research-data-center-data-access.html

Value of the data

- The data shows the overall distribution of SES and health for old people in Europe, and therefore allows us not only to disentangle how SES and health are related from childhood to old age, but also to examine the trajectories or dynamics of SES and health over the life course.
- SES information is based on various measurable variables to reflect multiple dimensions.
- Health data is also based on various variables and offers both subjective and objective measurements.
- The data enables an assessment of the notion of the long arm of childhood and is useful for life course analysis on health inequality [1].

1. Data

The data is based on SHARELIFE, i.e. the third wave (2008/2009) of the Survey of Health Ageing and Retirement in Europe (SHARE), which is a household panel survey [3]. SHARE is a cross-national panel database of micro data on health, socio-economic status and social and family networks of more than 45,000 individuals aged 50 or over in Europe and respondents are interviewed biennially. It is representative for the non-institutionalised population in European countries. The first wave started in 2004 and 12 countries participated, and in the sixth wave in 2015, 18 countries took part (http://www.share-project.org/home0/overview.html). SHARE is harmonised with its sister studies: the US Health and Retirement Study (HRS) and the English Longitudinal Study of Ageing (ELSA). In our sample, as inclusion criteria, the respondents are those aged 50 and above and the necessary variables needed for our research questions are available and comparable across countries, which eventually captures the citizens of 13 European countries. The data focuses on people's life histories, starting in childhood right through to old age. We use the data to try to understand how childhood experiences are associated with health in old age. In addition, we also take into account how mid-life conditions – such as education, income and occupation, and behaviour risk – mediate the effect of childhood SES and health on old age health.

Table 1
Sample size by country and by gender.

Country	Male	Female	Total
Austria	276	340	616
Germany	668	639	1307
Sweden	684	798	1482
The Netherlands	738	769	1507
Spain	678	521	1199
Italy	919	678	1597
France	714	845	1559
Denmark	647	726	1373
Greece	920	532	1452
Switzerland	415	462	877
Belgium	908	892	1800
Czech	589	740	1329
Poland	564	568	1132
Total	**8720**	**8510**	**17,230**

2. Experimental design, materials and methods

2.1. Data source

SHARELIFE is the third wave of SHARE which focuses on people's life histories and more than 30,000 men and women took part in this round of the survey. The respondents are representative for the population aged 50+. The retrospective part of SHARELIFE offers a complete history of changes in SES and health, beginning at childhood. It complements the SHARE panel data by providing life history information to enhance the understanding of how early life experiences – and events throughout life – influence the circumstances of older people. Thirteen European countries are included: Austria, Germany, Sweden, the Netherlands, Spain, Italy, France, Denmark, Greece, Switzerland, Belgium, Czech Republic and Poland. The distributions of our sample, separated into male and female, are presented in Table 1. In Table 2 we present the distribution of all variables used in our dataset.

2.2. Data description

See Tables 1 and 2.

2.3. Data design and measurements

The life course data is divided into three age groups to represent the longitudinal feature of the life cycle: childhood, adulthood, and old age. In childhood we have two latent variables, childhood SES and childhood health. The childhood SES (CH_SES), as one of the key independent variables, has three indicators or observed variables: (a) number of books in the household, which represents the cultural background and parents' education, (b) rooms per capita, which is a proxy for long-term household wealth, and (c) father's occupation, which we group into four categories according to ISCO (International Standard Classification of Occupation) skills levels: elementary occupations, skilled (service, shop or market sales worker, skilled agricultural or fishery worker, craft or related trades worker, and plant/machine operator or assembler), associate (technician or associate professional, clerk), and manager (legislator, senior official or manager, professional). All SES indicators refer to when the respondents were aged 10. The second latent variable is childhood health (CH_Health), which is constructed using three indicators: (a) childhood self-rated health, a five-point scale of health, from poor to excellent, (b) a binary variable indicating if the individual ever missed school for at least one month because of health, and (c) a binary variable indicating if the individual was ever hospitalised for at least one month. These health indicators refer to respondents aged up to 15 years old.

Table 2
Descriptive statistics of the data set. Percentages are shown for categorical variables; actual numbers of cases are provided in parentheses. Means, standard deviation and ranges are shown for continuous variables.

Variables	Category	All (N=17,230)
Age	Mean	70.47
	Std. Deviation	7.62
Number of books	Up to 10 books (=1)	46.12 (7866)
	11–25 books	22.16 (3780)
	26–100 books	19.79 (3375)
	101–200 books	6.06 (1033)
	> 200 books (=5)	5.88 (1003)
	Missing	*173*
Rooms per capita	Mean	0.71
	Std. Deviation	0.42
	Range	0: 8.75
	Missing	*213*
Father's occupation	Elementary (=1)	18.09 (2941)
	Skilled	69.42 (11,287)
	Associate	4.22 (686)
	Manager (=4)	8.27 (1345)
	Missing	*971*
Childhood SRH	Excellent (=5)	34.05 (5836)
	Very good	33.69 (5775)
	Good	23.92 (4100)
	Fair	6.15 (1055)
	Poor (=1)	2.19 (375)
	Missing	*89*
Ever missed school	No (=1)	88.84 (15,250)
	Yes (=0)	11.16 (1916)
	Missing	*64*
Ever in hospital	No (=0)	94.05 (16,172)
	Yes (=1)	5.95 (1024)
	Missing	*34*
Education (years of schooling)	Mean	10.34
	Std. Deviation	4.31
	Range	0: 25
	Missing	*2115*
Occupation	Elementary (=1)	19.40 (3280)
	Skilled	57.09 (9654)
	Associate	8.78 (1484)
	Manager (=4)	14.74 (2493)
	Missing	*319*
Income	Mean	16,779.89
	Std. Deviation	17,720.65
	Range	0: 586047.1
	Missing	*1916*
Wealth	Mean	141,249.20
	Std. Deviation	227,077.80
	Range	− 341,522.3: 6,932,346
	Missing	*2094*
Smoking	Current (=1)	16.46 (2715)
	Former	31.57 (5207)
	Never (=3)	51.97 (8573)
	Missing	*735*

Table 2 (*continued*)

Variables	Category	All ($N=17,230$)
Physical activity	Non-active (=1)	11.05 (1828)
	Active (=2)	88.95 (14,716)
	Missing	*686*
Old age health	Poor (=1)	14.18 (2441)
	Fair	29.29 (5042)
	Good	36.34 (6256)
	Very good	13.95 (2402)
	Excellent (=5)	6.24 (1075)
	Missing	*14*

The second stage is adulthood. First, education is measured by years of schooling. The second covariate is a latent variable, adult SES, and we construct it using the indicators occupation (according to ISCO), household income and wealth – where income and wealth are corrected by purchasing power parities (PPP), relative to German Euros in 2006. Adulthood indicators refer to respondents who are between 30 and 60 years old.

The third stage contains the variables in old age, i.e. behavioural risks and old age health. We consider smoking (currently smoking, former smoker, and never having smoked) and physical activities (non-active and active) which require moderate level of energy such as gardening, cleaning the car, or taking a walk. Old age health is based on the question "Would you say your health now is…", with the possible responses: "poor", "fair", "good", "very good", and "excellent". The old age variables refer to respondents aged 60 and above.

3. Method

With this data it is possible to employ a structural equation model (SEM) approach, which allows us to test direct and mediating (indirect) effects via path analysis, and to combine this with measurement models for SES and health to reduce measurement error [4–8]. Estimating direct, indirect and total effects in our model represents the two ways in which childhood exerts its influence (the long arm) on health in old age. By direct effect we mean the extent to which childhood SES and childhood health affect old age health directly – that is, unmediated by any other variables – whereas the total effect is the sum of the direct and the indirect effects (i.e., those mediated by at least one intervening variable). Data preparation is performed using Stata 14.1, including the newspell package [2].

Acknowledgements

The authors acknowledge the financial support from European Research Council [Grant number: 313532].

This paper uses data from SHARE wave 4 release 1.1.1, as of March 28th 2013 (DOI: 10.6103/SHARE.w4.111) or SHARE waves 1 and 2 release 2.6.0, as of November 29th 2013 (DOIs: 10.6103/SHARE.w1.260 and 10.6103/SHARE.w2.260) or SHARELIFE release 1.0.0, as of November 24th 2010 (DOI: 10.6103/SHARE.w3.100). The SHARE data collection has been primarily funded by European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006–062193, COMPARE, CIT5- CT-2005–028857, and SHARELIFE, CIT4-CT-2006–028812) and through the 7th Framework Programme (SHARE-PREP, No. 211909, SHARE-LEAP, No. 227822 and SHARE M4, No. 261982). Additional funding from U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA

04–064) and German Ministry of Education and Research, as well as from various national sources, is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at: http://dx.doi.org/10.1016/j.dib.2016.11.094.

References

[1] E. Pakpahan, R. Hoffmann, H. Kröger, The long arm of childhood circumstances on health in old age: evidence from SHARELIFE, Adv. Life Course Res. (2016), http://dx.doi.org/10.1016/j.alcr.2016.10.003.
[2] H. Kröger, Newspell: easy management of complex spell data, Stata J. 15 (2015) 155–172.
[3] A. Börsch-Supan, M. Brandt, C. Hunkler, T. Kneip, J. Korbmacher, F. Malter, B. Schaan, S. Stuck, S. Zuber, Data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE), Int. J. Epidemiol. (2013) dyt088. http://dx.doi.org/10.1093/ije/dyt088.
[4] M.S. Mulatu, C. Schooler, Causal connections between socio-economic status and health: reciprocal effects and mediating mechanisms, J. Health Soc. Behav. 43 (2002) 22.
[5] J.R. Warren, Socioeconomic status and health across the life course: a test of the social causation and health selection hypotheses, Soc. Forces 87 (2009) 2125–2153. http://dx.doi.org/10.1353/sof.0.0219.
[6] K.A. Bollen, Structural Equations with Latent Variables, 1 edition,. Wiley-Interscience, New York, NY, 1989.
[7] E. Pakpahan, R. Hoffmann, H. Kröger, Statistical methods for causal analysis in life course research: an illustration of a cross-lagged structural equation model, a latent growth model, and an autoregressive latent trajectories model, Int. J. Soc. Res. Methodol. 0 (2015) 1–19. http://dx.doi.org/10.1080/13645579.2015.1091641.
[8] J. Wang, X. Wang, Structural Equation Modeling: Applications Using Mplus, 1 edition,. Wiley, Chichester, West Sussex England; Hoboken, N.J, 2012.